

Intra-examiner repeatability and agreement in accommodative response measurements

B Antona, I Sanchez, A Barrio, F Barra and E Gonzalez

Departamento de Optica II (Optometria y Vision), Escuela Universitaria de Óptica, Universidad Complutense de Madrid, C/ Arcos de Jalón s/n, 28037 Madrid, Spain

Abstract

Purpose: Clinical measurement of the accommodative response (AR) identifies the focusing plane of a subject with respect to the accommodative target. To establish whether a significant change in AR has occurred, it is important to determine the repeatability of this measurement. This study had two aims: first, to determine the intraexaminer repeatability of AR measurements using four clinical methods: Nott retinoscopy, monocular estimate method (MEM) retinoscopy, binocular crossed cylinder test (BCC) and near autorefractometry. Second, to study the level of agreement between AR measurements obtained with the different methods.

Methods: The AR of the right eye at one accommodative demand of 2.50 D (40 cm) was measured on two separate occasions in 61 visually normal subjects of mean age 19.7 years (range 18–32 years). The intraexaminer repeatability of the tests, and agreement between them, were estimated by the Bland–Altman method. We determined mean differences (MD) and the 95% limits of agreement [coefficient of repeatability (COR) and coefficient of agreement (COA)].

Results: Nott retinoscopy and BCC offered the best repeatability, showing the lowest MD and narrowest 95% interval of agreement (Nott: -0.10 ± 0.66 D, BCC: -0.05 ± 0.75 D). The 95% limits of agreement for the four techniques were similar (COA = ± 0.92 to ± 1.00 D) yet clinically significant, according to the expected values of the AR. The two dynamic retinoscopy techniques (Nott and MEM) had a better agreement (COA = ± 0.64 D) although this COA must be interpreted in the context of the low MEM repeatability (COR = ± 0.98 D).

Conclusions: The best method of assessing AR was Nott retinoscopy. The BCC technique was also repeatable, and both are recommended as suitable methods for clinical use. Despite better agreement between MEM and Nott, agreement among the remaining methods was poor such that their interchangeable use in clinical practice is not recommended.

Keywords: accommodative response, agreement, dynamic retinoscopy, measurement, repeatability

Introduction

Through accommodation, the dioptric power of the eye increases so that images of close objects can be brought into focus on the retina. In an optometric examination,

accommodative function can be assessed by measuring the amplitude of accommodation, accommodative response and accommodative facility (Saladin, 1998). If these three factors are not measured, an accommodative dysfunction could pass unnoticed (Wick and Hall, 1987; Goss, 1992). Measuring the accommodative response (AR) establishes the subject's plane of focus with respect to the accommodative target; that is, whether there is over- or under-accommodation. In clinical practice, rather than using the term accommodative response, it is common to consider the error of accommodation, which is the difference between the accommodative stimulus and the accommodative response. We use the term accommodative lag if this

Received: 5 June 2008

Revised form: 28 October 2008; 17 February 2009; 26 March 2009; 23 April 2009

Accepted: 10 May 2009

Correspondence and reprint requests to: Beatriz Antona.

Tel.: +34 91 394 68 94; Fax: +34 91 394 68 85.

E-mail address: bantona@opt.ucm.es

1 difference is positive (under-accommodation) and
2 accommodative lead if the difference is negative (over-
3 accommodation).

4 As a consequence of increasing near vision demands
5 in modern society, accommodation problems are a
6 common cause of ocular asthenopia (Rosenfield and
7 Gilmartin, 1990). To diagnose and treat several accom-
8 modative and vergence disorders, it is important to
9 assess the AR to a near target. Under-accommodation
10 or over-accommodation with respect to the object plane
11 is a frequent cause of asthenopia (Birnbaum, 1993). A
12 low AR, that is, an accommodative lag $\geq +1.00$ D,
13 could be the outcome of uncorrected or undercorrected
14 hypermetropia, presbyopia, or an accommodative defi-
15 ciency. However, an accommodative lead is mainly
16 associated with a latent hyperopia, an accommodation
17 excess (pseudomyopia) or with an accommodative
18 spasm (Elliott, 2007). Similarly, a significant difference
19 between the AR of each eye could indicate refractive
20 and/or accommodative disequilibrium (Saladin, 1998).
21 On the other hand, a reduced ability to accommodate
22 for close targets has been implicated in the development
23 of myopia and it is known that adults with myopia of
24 recent onset accommodate significantly less than emme-
25 tropic subjects when viewing close targets (Gwiazda
26 *et al.*, 1993).

27 Several objective and subjective methods exist to
28 measure the AR. The most common techniques are
29 forms of dynamic retinoscopy, the most widely known
30 being the monocular estimate method (MEM retinos-
31 copy) and Nott retinoscopy (Locke and Somers, 1989;
32 Rosenfield *et al.*, 1996; Cacho *et al.*, 1999; McClelland
33 and Saunders, 2003; Goss *et al.*, 2005). A common
34 subjective technique is the binocular crossed cylinder
35 test (BCC), whilst the open-field autorefractometer
36 technique is a valid and repeatable objective method
37 of measuring the AR (Chat and Edwards, 2001; Mallen
38 *et al.*, 2001).

39 In general, it is considered that MEM retinoscopy
40 provides a good measure of AR (Rouse *et al.*, 1982) and
41 the method shows good inter-examiner repeatability
42 (Goss *et al.*, 2005). More repeatable measures (espe-
43 cially in young adults) might be expected using tests that
44 do not require the insertion of additional lenses to assess
45 the AR, e.g., Nott retinoscopy (Cacho *et al.*, 1999).

46 The BCC test is an easy clinical method to determine
47 the AR, but it seems to have a series of theoretical
48 shortcomings. The use of linear targets does not help in
49 placing the circle of least confusion on the retina (which
50 is the basis of the technique) due to their horizontal and
51 vertical orientation. It is therefore arguable whether
52 subjects will prefer to keep both groups of lines blurred
53 by placing the circle of least confusion on the retina or
54 would prefer to modify their AR to improve the clarity
55 of one of the groups of lines. In this latter case, the

accommodative error measured would not reflect the
habitual response of the patient to a more natural
stimulus (Rosenfield *et al.*, 1996). In current practice,
this test is used more to determine the tentative add for
near vision than to assess the AR (Elliott, 2007).

Among the optometric methods available for deter-
mining the AR, most clinicians select one procedure for
routine use depending on their personal preference.
However, to determine whether a significant change has
occurred in a clinical variable, it is essential to know the
repeatability of the measurement procedure. Several
studies have compared the different techniques for
measuring AR (Locke and Somers, 1989; Rosenfield
et al., 1996; Cacho *et al.*, 1999; McClelland and Saun-
ders, 2003; Goss *et al.*, 2005), but very few have assessed
their repeatability: McClelland and Saunders (2003)
only studied the repeatability of AR measurements
obtained by Nott retinoscopy and by open-field auto-
refractor at three unusual near distances of 10, 16.7 and
25 cm. McKee (1981) and Goss *et al.* (2005) studied the
inter-examiner repeatability of Nott and MEM.
The present study was therefore designed to establish
the intra-examiner repeatability and agreement of mea-
surements made using four different methods to deter-
mine the AR in clinical practice: (1) Nott retinoscopy,
(2) MEM retinoscopy, (3) the binocular crossed cylinder
test, (4) and open-field autorefraction measures at near.

Methods

Study population

The population comprised 61 subjects aged 18–32 years
(mean 19.74, S.D. 2.5 years) recruited following infor-
mative talks from the first year students of the School of
Optics, Universidad Complutense de Madrid, Spain.
This population was selected on the grounds that having
recently been admitted, the subjects were unaccustomed
to the type of tests performed. They can therefore be
considered representative of a typical clinical population
of this age range with similar near work demands. The
study design fulfilled the tenets of the Declaration of
Helsinki and all the subjects signed an informed consent
form to participate in the study. The research was
approved by the Ramon and Cajal Hospital of Madrid
Ethical Committee.

The subjects were first required to complete a ques-
tionnaire to record their age, sex, ocular and medical
history and several questions to determine whether the
subject had binocular vision symptoms. They had to mark
with a cross if they experienced any of the following
symptoms: (1) eyestrain, (2) occasional double vision, (3)
blurred vision after close work or (4) headaches related to
the use of the eyes for near tasks. We excluded from the
study those subjects who had one or more of these

symptoms. Next, the optometric characteristics of each subject were determined by several screening tests. The clinical criteria for inclusion were:

- (1) A corrected visual acuity (VA) ≥ 0.9 decimal visual acuity of Snellen (20/22) in each eye at distance and near.
- (2) No ocular pathology.
- (3) No history of refractive surgery, strabismus, nystagmus or amblyopia.
- (4) No medication or disease that could affect accommodation, fusional vergences or ocular motility.
- (5) No accommodative or vergence anomalies. Binocular test results were compared with the cut-off values of Table 1, derived from the normal values of Scheiman and Wick (2008). When the subject had one or more of the parameters outside the cut-off values, he or she was eliminated from the study.

According to Bland and Altman (1986), the best way to assess the repeatability of a technique is to take several measurements in a series of subjects. Thus, measurements were taken on two separate occasions separated by a time interval of at least 24 h and, at most, 10 days. This retest interval is short enough to minimize the probability of encountering real changes in visual abilities and at the same time sufficiently long to reduce the chances of a significant learning effect (Argimon and Jiménez, 2004).

There were four examiners, one for each technique. Each test was administered by the same examiner who was masked to the AR measurements obtained using the other methods. Each examiner took only one measurement at each test session. The tests in the two sessions were undertaken by the same examiner to determine intra-examiner repeatability. The examiner was masked to the first set of measurements during the second session. Each test was performed in exactly the same manner in each subject and at approximately the same time of day to reduce the effect of diurnal variation on the variables examined (Levine *et al.*, 1985). The second set of tests was similarly undertaken.

AR assessment

The accommodative response was assessed by: (1) Nott retinoscopy, (2) MEM retinoscopy, (3) binocular crossed cylinders and (4) near open-field autorefraction measures. Although two of these tests (1 and 4) provide a direct measure of the AR, results are expressed as accommodative error (that is, in terms of a lag or a lead of accommodation) for all the procedures (Zadnik, 1997). Since the accommodative demand used was always 2.50 D, the accommodative error was calculated as 2.50 D minus the accommodative response. The MEM retinoscopy and BCC techniques directly provide accommodative errors such that no conversion was required.

Nott retinoscopy was always performed first to avoid the possible effects of the plus and minus lenses on the AR (Cacho *et al.*, 1999; Goss *et al.*, 2005), while the order of the other tests was randomly established to avoid a learning effect and/or subject fatigue influencing the results. Time between techniques was at least 5 min. Each subject was requested to pick out a numbered ball to indicate the order of the tests.

Although measurements were only recorded for the right eye, the tests were conducted binocularly so as not to eliminate the effect of convergence accommodation (Rosenfield *et al.*, 1996). Subjects wore the full far-distance subjective refractive correction, and the target was illuminated directly in a similar way for all the methods to provide a good control of the AR. The protocols followed for each technique are described below.

Nott retinoscopy. The subject remained seated and looking through the phoropter. The target was illuminated so that room lights may be dimmed, since young subjects tend to over-accommodate when the test card is poorly illuminated (Locke and Somers, 1989). As a fixation target, a narrow near-point card was placed at 40 cm on the phoropter's near vision rod, aligning the

Table 1. Cut-off values for the screening tests

Test	Method	Cut-off values for inclusion	
Amplitude of accommodation	Push-up test	≥ 6 D	
Binocular accommodative facility	Flipper ± 2.00	≥ 5 cpm	
Near point of convergence	Accommodative test	Break point ≤ 7.5 cm	
		Recovery point ≤ 10 cm	
Step vergence testing	Prism bar	Far vision	PFV $\geq 4/5 \Delta$
			NFV $\geq 4/2 \Delta$
		Near vision	PFV $\geq 10/7 \Delta$
			NFV $\geq 7/5 \Delta$
Vergence facility test	3 Δ BI/12 Δ BO	≥ 15 cpm	
Stereoacuity	Randot test	20 s arc	

BI, base in; BO, base out; PFV, positive fusional vergence; NFV, negative fusional vergence.

zero with the corneal apex. The card had several lines of letters of high contrast ($\approx 80\%$) black on white. The subject was instructed to view with both eyes the line of letters corresponding to a VA equivalent to a Snellen value of 1.0 (6/6). To ensure that the technique measures the accommodative response throughout the duration of the test, the subject was asked to keep the letters as clear as possible during testing (Saladin, 1998).

Initially the retinoscope was placed 40 cm from the subject alongside the target to observe the horizontal meridian of the right eye. If the subject was accommodating accurately to the target, then the reflex seen alongside the target will be neutral. If the initial reflex was 'against', the examiner moved the retinoscope towards the subject to find neutral. If the initial reflex observed was 'with', the retinoscope should be moved away from the subject to achieve neutrality. The linear distance from the corneal apex to the position of the retinoscope at neutrality was recorded to the nearest ± 0.5 cm and converted to dioptres. The accommodative error was obtained by subtracting the measured AR from the target stimulus at 40 cm (2.50 D).

MEM retinoscopy. The tests were performed with the subject's distance correction on a trial frame. The subject fixated on an illuminated target with the room lights turned down as for the Nott procedure. The fixation target was a near vision card we made attached to the retinoscope's window with several lines of letters at high contrast ($\approx 80\%$) black on white.

Initially, the retinoscope was placed 40 cm from the spectacle plane of the subject measured using a tape. The subject was instructed to view with both eyes the line of letters corresponding to a Snellen decimal VA of 1.00, and to try to keep the letters clear. The reflex was observed at the horizontal meridian of the right eye. The examiner introduced lenses (steps of 0.25 D) in front of the subject's eye until neutrality was first observed (Locke and Somers, 1989; Rosenfield *et al.*, 1996; Cacho *et al.*, 1999; McClelland and Saunders, 2003; Goss *et al.*, 2005). The lenses were inserted and removed quickly (< 0.5 s) to avoid changing the subject's accommodative status (Campbell *et al.*, 1998). The lenses used directly measured the accommodative error; a minus lens indicated over-accommodation and a plus lens indicated under-accommodation.

Binocular crossed cylinders (BCC). The fixation target was a card at 40 cm with a grid formed by black-on-white horizontal and vertical lines at high contrast ($\approx 80\%$). Room and column lights were kept on. The phoropter wheel's stationary crossed-cylinder lens was inserted before each eye. This lens corresponded to crossed cylinders of ± 0.50 D with the vertical axis at 90° . The subject was then asked whether the horizontal

or vertical lines appear clearest. If both could be seen equally clearly, the result of the test (accommodative error) was recorded as 0.00 D.

At the starting situation described, it is common for the subject to view the horizontal lines clearer due to under-accommodation. In this case, plus lenses were added in 0.25 diopter steps until the subject viewed the horizontal and vertical lines equally clear or if this did not occur, the examiner recorded the first lens resulting in the vertical lines appearing clearer. In contrast, if initially the vertical lines appeared clearer, this meant the subject was over-accommodating and minus lenses were added until they could both be clearly seen.

Autorefractometry. The autorefractometer Shin-Nippon SRW-5000, also known as the Grand Seiko WV-500 autorefractor (Grand Seiko Company, Fukijama, Japan), is an open-field infrared binocular autorefractometer that allows the subject to view a real card at any distance such that the AR can be assessed. Previous studies have determined it provides valid, reproducible measures both in adults and children (Chat and Edwards, 2001; Mallen *et al.*, 2001) and AR measures can be obtained using external cards (Wolffsohn *et al.*, 2001, 2004; Nakasuka *et al.*, 2003, 2005).

To measure the AR, the external card was placed in a fixed support 40 cm from the subject's corneal apex. The subject was asked to make every effort to keep the letters on the card clear: these corresponded to a VA equivalent to a Snellen value of 1.0 (6/6). Three autorefraction measures for the right eye were averaged and the mean spherical equivalent obtained. To calculate the accommodative error, the accommodative demand applied (2.50 D) was added to the mean spherical equivalent such that a negative result indicated over-accommodation and a positive result under-accommodation.

Statistical analysis

Once the data for the whole sample had been collected they were analysed using the Analyse-it for Microsoft Excel program (Leeds, UK. See <http://www.analyse-it.com>) and SPSS v. 11 for Windows software (SPSS Inc., Chicago, IL, USA).

The Bland–Altman method was used to determine the degree of repeatability of the tests and agreement among them (Bland and Altman, 1986; Zadnik *et al.*, 1992). From a clinical perspective, the advantage of this method is that the agreement of the tests is expressed in the same units of measurement as the test itself and this allows the clinician to establish his own criteria as to whether or not a difference is significant. This method was used when the differences, as established by the Anderson–Darling normality test (D'augostino and Stevens, 1986), showed a normal distribution.

The factors determined were the mean difference (MD), the standard deviation (S.D.), the coefficient of repeatability (COR = $1.96 \times \text{S.D.}$) and the limits of agreement at the 95% level (mean difference \pm COR). When the normality test revealed a non-normal distribution, we determined the 95th percentile of the absolute values of the differences, instead of calculating the COR. Similarly, we determined the coefficient of agreement (COA) among the tests. Finally, we also conducted a two-factor (method and session) repeated measures ANOVA.

Given the sample size, a small difference could be statistically significant yet not clinically significant. Differences from the mean were plotted to establish the limits of agreement at the 95% level and to estimate the repeatability of the measures. The limit of agreement constitutes a threshold for the differences in successive measures that has to be surpassed if the difference indicates that a change in the value has really occurred and cannot simply be explained by natural variation among measurements.

Results

Seventy-two participants came to the first appointment. The review of the completed questionnaire led us to

exclude six subjects: one with a history of refractive surgery; two with strabismus; one woman with multiple sclerosis; and two symptomatic subjects. Five more subjects were eliminated on the basis of the results of the preliminary tests: two with VA lower than 0.9; two with a too remote NPC; and one with a low value of stereoacuity. So the final sample was composed of 61 subjects.

Repeatability

Figure 1 shows difference vs mean plots (Bland–Altman) with the difference between two measures (Final – Initial) on the y-axis plotted against the average of the two measures on the x-axis. These plots are useful for comparing the amplitudes of intervals of concordance at 95% for the different methods. For each plot, if measures show good repeatability, the averaged difference will be close to zero and the ± 1.96 S.D., or 95% limits of agreement, will be small. By direct observation, none of the plots revealed a clear tendency for the difference to increase with the dioptric value, i.e. the repeatability of the tests was independent of the accommodative response.

Table 2 shows the repeatability of the results obtained for the four methods of measuring the AR. Nott

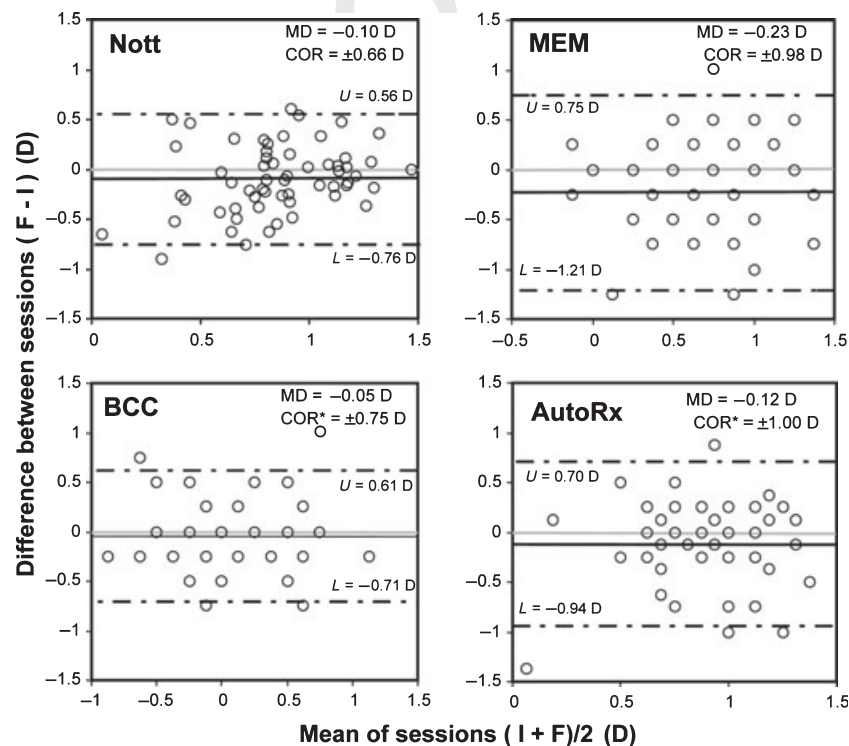


Figure 1. Mean differences (MD), upper (U) and lower (L) Bland–Altman 95% agreement limits, and coefficients of repeatability (COR) between Final (F) and Initial (I) accommodative response (AR) measures by Nott retinoscopy, monocular estimate method (MEM) retinoscopy, the binocular crossed cylinder (BCC) method, and near autorefractometry (AutoRx). The BCC and AutoRx measures were non-normally distributed and the COR was taken as the 95th percentile of the absolute value of the differences (designated COR*).

Table 2. Repeatability in accommodative response measures

	Accommodative error	MD (Final – Initial)	COR (D)
	Mean ± S.D. (D)	± S.D. (D)	
Nott retinoscopy	0.83 ± 0.34	–0.10 ± 0.33	±0.66
MEM retinoscopy	0.63 ± 0.50	–0.23 ± 0.50	±0.98
BCC	0.10 ± 0.38	–0.05 ± 0.34	±0.75*
Near autorefractometry	0.95 ± 0.40	–0.12 ± 0.42	±1.00*

S.D., standard deviation; MD, mean difference; COR, coefficient of repeatability; MEM, monocular estimate method; BCC, binocular crossed cylinders.

*When COR appears with an asterisk, the value given is the 95th percentile of the absolute difference.

retinoscopy showed the highest repeatability, since the 95% agreement interval was the smallest (COR = ±0.66 D) and the mean difference was low (MD = –0.10 D).

Agreement

Table 3 shows the results of our analysis of agreement among the different techniques used to determine the AR. In each case, agreement intervals were similar although clinically high (around ±1.00 D) because agreement coefficients were high, compared to the expected values of the AR (Scheiman and Wick, 2008), with the single exception of better agreement between the two dynamic retinoscopy techniques (COA = ±0.64 D). However, this better COA should be considered with respect to the large COR of the two dynamic retinoscopy techniques (COR: Nott = ±0.66 D, MEM = ±0.98 D). Because the agreement (COA = ±0.64 D) was better than the repeatability of MEM retinoscopy (COR = ±0.98 D) the agreement must be interpreted as lower than the COA suggested.

Bland and Altman (1986) suggested that rather than determining correlation coefficients, a more informative

Table 3. Agreement between tests measuring the accommodative response

	MD ± S.D. (D)	COA (D)
Nott – MEM	+0.23 ± 0.32 (Nott > MEM)	±0.64
Nott – BCC	+0.72 ± 0.48 (Nott > BCC)	±0.95
Nott – AutoRx	–0.13 ± 0.50 (AutoRx > Nott)	±0.98*
MEM – BCC	+0.52 ± 0.51 (MEM > BCC)	±1.00
MEM – AutoRx	–0.31 ± 0.50 (AutoRx > MEM)	±0.98
BCC – AutoRx	–0.82 ± 0.46 (AutoRx > BCC)	±0.92

MD, mean difference; COA, coefficient of agreement; Nott, Nott retinoscopy; MEM, monocular estimate method; AutoRx, near autorefractometry; BCC, binocular crossed cylinders.

*When COA appears with an asterisk, the value given is the 95th percentile of the absolute difference.

method of comparing the results of two methods of testing is to plot the differences between the results obtained by the two methods (along the y-axis) against the mean of the two methods (along the x-axis). Horizontal lines are plotted to indicate the mean of the differences (MD) and the limits of agreement between the two methods: i.e., MD ± 1.96 S.D. of these differences. Such plots for our AR data are shown in Figure 2. Since the scale for the y-axis is the same in all the plots, it can be easily seen that no marked differences exist in agreement among the accommodative error values provided by the different methods with the exception of the two dynamic retinoscopy methods compared, which yielded a narrower agreement interval indicating better agreement between the two techniques. None of the plots revealed a tendency for the difference between methods to increase as the dioptric value increased. Thus, the level of inter-method agreement is independent of the accommodative response.

As a complement to the Bland–Altman method to study the repeatability and the agreement between techniques, we also conducted a two-factor (method and session) repeated measures ANOVA. This method revealed significant differences among the tests ($F_{3,180} = 67.31$; $p < 0.0001$) and between the two measurement sessions ($F_{1,6} = 19.26$; $p < 0.0001$). However, no interaction effects between the two factors were detected ($F_{3,180} = 2.53$; $p = 0.06$). A post-hoc analysis using the Scheffé technique indicated significant differences between each of the mean values ($p < 0.05$) except for Nott retinoscopy vs autorefractometry ($p = 0.33$) or MEM retinoscopy ($p = 0.06$). Table 2 shows the mean ARs recorded using the different procedures. MEM and Nott tended to provide higher accommodative lags (lower accommodative response) than the BCC method. Open-field autorefractometry was the technique that revealed the largest accommodative lags.

Discussion

Repeatability

One of the objectives of this study was to determine the repeatability of four techniques used by clinicians to assess the accommodative response. Nott retinoscopy appeared as the most repeatable, for which the differences between measures of up to ±0.66 D can be attributed to measurement errors. In contrast, the techniques emerging as least repeatable were MEM retinoscopy and near open-field autorefractometry, for which the measurement error can be as high as ±1.00 D (see Table 2). Whitefoot and Charman (1992) suggested that normal variations in tonic accommodation and the near phoria effect could contribute to this wide variability in AR measures. They

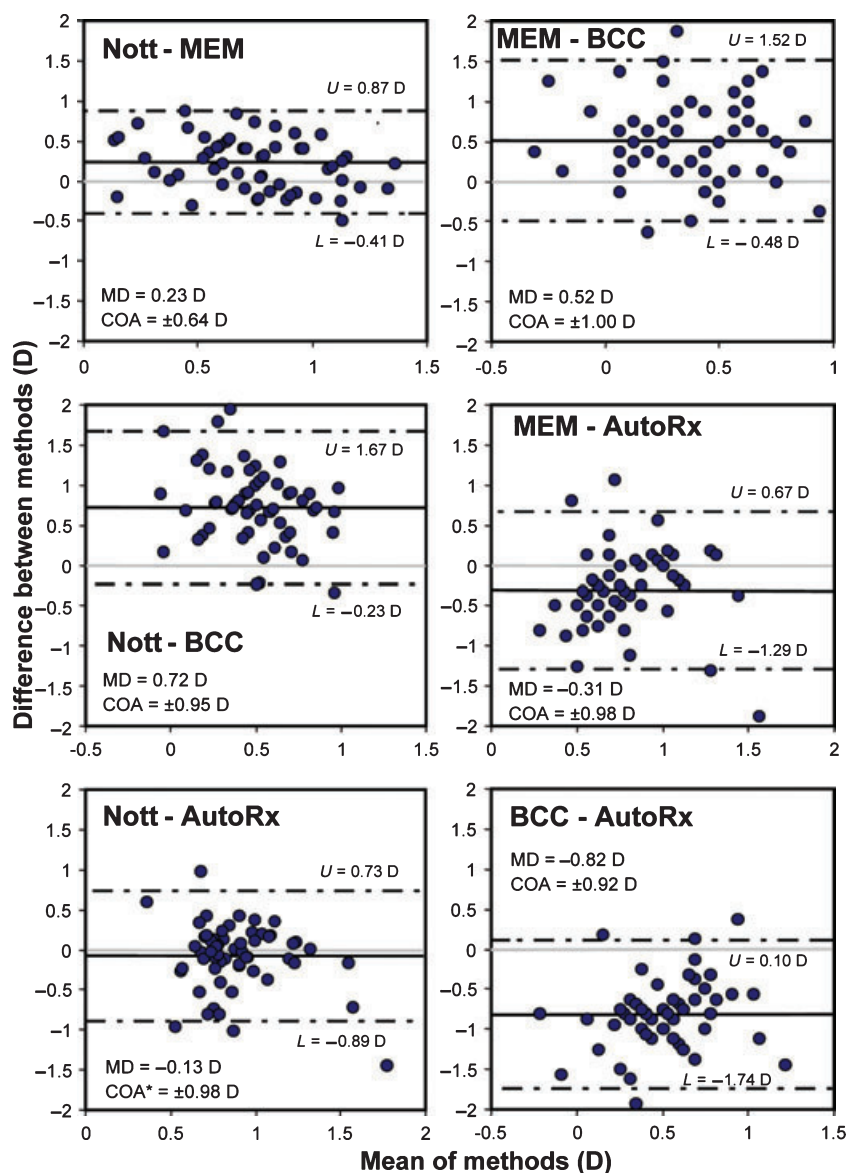


Figure 2. Bland–Altman assessment of agreement between the four different methods used to determine the accommodative response. The continuous line represents the mean difference (MD) between the measures obtained with the different techniques. Broken lines indicate the upper (U) and lower (L) limits of the 95% agreement interval ($MD \pm 1.96 \times S.D.$). When differences showed a non-normal distribution, the 95th percentile of the absolute value of the differences was taken as the coefficient of agreement (COA^*).

reported that there is a tendency for the lag of accommodation in dynamic retinoscopy to be higher in esophoria than exophoria. This is probably the result of a greater convergence effort in exophoria which, by stimulating more convergence accommodation, tends to reduce the lag of accommodation (higher AR) (Whitefoot and Charman, 1992).

When the AR is assessed by autorefractometry in an ametropic individual, the accommodative response has to be measured through lenses that compensate the refractive error. However, none of the marketed autorefractometers was designed with this in mind since the

power and position of these compensating lenses are not taken into account in the autorefractometer's calculations. These non-considered variables could represent a systematic error in autorefractometry measures of AR possibly contributing to the low repeatability detected in the present study (Kimura *et al.*, 2007).

McClelland and Saunders (2003) assessed the repeatability of Nott retinoscopy in estimating the AR of a population of subjects aged 6–43 years (mean: 24.45, S.D.: 9.82 years). As in the present study, Nott retinoscopy appeared as a repeatable technique, since no significant differences were observed between two sets of

measures taken for the different accommodative demands tested in the study (10, 6 and 4 D). Coefficients of repeatability were ± 1.34 D for a 10 D demand, ± 1.09 D for 6 D and ± 0.56 D for 4 D.

A possible limitation of our study is that Nott was always performed first, instead of its order being randomized with the other techniques. However, this was done to eliminate the possible effects of lenses on the AR and to compare our results to those of others (Cacho *et al.*, 1999; Goss *et al.*, 2005). Another possible limitation is that we only took a single measure for each technique in each session. We followed that protocol to minimize learning effects and because it is a situation more similar to that in clinical practice.

Agreement

Rosenfield *et al.* (1996) evaluated the AR for a 2.50 D target stimulus in a series of 24 subjects aged 25.5–30.1 years (mean: 25.0; S.D.: 0.4 years). They used the Canon Autoref R-1 autorefractometer; two dynamic retinoscopy techniques (neutralization by lenses and by varying the work distance); and three subjective methods (near bichromatic test and crossed cylinders at near without myopization or under a myopization of +1.00 D over the far subjective refraction of the subject). When they compared the autorefractometer results with those provided by the other tests, best agreement was found with the findings of dynamic retinoscopy, obtaining neutrality by varying the working distance (COA = ± 0.65 D). Agreement between the remaining techniques and autorefractometry was worse or similar to that obtained in the present study, in which COAs were close to ± 1.00 D. In the study by Goss *et al.* (2005), two examiners measured the AR using three techniques (MEM retinoscopy, Nott retinoscopy and the Canon Autoref R-1 autorefractometer) in 50 subjects (aged 20–35, mean 24.8 years). The mean differences between MEM and autorefractometry (MD = 0.51 D) and between Nott and autorefractometry (MD = 0.51 D) were greater than those found in the present study (Nott vs AutoRx: MD = -0.13 D; MEM vs AutoRx: MD = 0.31 D). However, their coefficients of agreement (Nott vs AutoRx: COA = ± 0.59 D; MEM vs AutoRx: COA = ± 0.82 D) were lower than those reported here of ± 0.98 D in both cases. In the study by Goss *et al.* (2005), however, since the values obtained by two examiners were averaged, extreme values may have been eliminated to give a narrower agreement interval.

McClelland and Saunders (2003) also compared the Nott AR results to those obtained using the Shin-Nippon SRW-5000 autorefractometer. Nott retinoscopy showed good agreement with those obtained objectively using the autorefractometer. Notwithstanding, for the demand of

4 D, Nott retinoscopy tended to give a lower accommodative lag than the autorefractometer (MD = 0.1 D), a behaviour also exhibited by our results (see Table 3).

In a study comparing the AR determined using BCC vs dynamic retinoscopy techniques, Locke and Somers (1989) observed significantly lower accommodative lags for the BCC test, a finding consistent with the results of the present study and with that of Rosenfield *et al.* (1996). Moreover, there seems to be no great influence of performing the test under discrete myopization (adding +1.00 D over the far subjective refraction) or without myopizing (Rosenfield *et al.*, 1996). In the BCC, the addition of lenses determines that the AR is not static in young adults, such that the accommodative lag cannot be correctly evaluated. For instance, a patient who initially reports seeing the horizontal lines clearest, may continue to give the same reply after the insertion of plus lenses and the consequent relaxation of accommodation. In this situation, the end point may be conditioned not only by the AR but also by the subject's ability to relax accommodation (Goodson and Afanador, 1974).

When Locke and Somers (1989) compared several dynamic retinoscopy techniques, no significant differences in ARs were obtained using MEM and Nott retinoscopy. Jackson and Goss (1991) also reported good agreement between these two procedures in a series of 244 school children. However, Cacho *et al.* (1999) assessed the AR determined first through Nott and then MEM retinoscopy in 50 subjects aged 15–35 years (mean: 23.96, S.D.: 3.17 years) and did find significant differences ($p < 0.05$) (MD = 0.41 D using MEM > Nott; COA = ± 0.59 D). Garcia and Cacho (2002) in another comparative assessment of the same techniques in subjects with binocular defects concluded that despite the techniques being closely correlated, the limits of agreement indicated the Nott method was the most recommended. Goss *et al.* (2005) reported, using the average for two examiners, good agreement between the two techniques with a COA of ± 0.55 D, which is similar to the ± 0.48 D found by Rosenfield *et al.* (1996). In the study performed here, although best agreement was between the two dynamic retinoscopy methods, the 95% interval of agreement was slightly wider (COA = ± 0.64 D).

In conclusion, agreement between the four AR techniques was poor, indicating that the methods are not interchangeable. We recommend that the practitioner should opt for the use of the most repeatable methods: Nott retinoscopy and BCC. The patient's clinical information should be assessed according to the mean differences (MD) observed in the repeatability study and the practitioner should then decide whether the diagnosis or treatment would differ in the context of the 95% limits of agreement. Considering

extreme values within the interval of agreement at the 95% level is especially important in patients in whom the signs and symptoms observed are inconsistent.

References

- Argimon, J. M. and Jiménez, J. (2004) *Métodos de Investigación Clínica y Epidemiológica*. Elsevier, Madrid, pp. 90–100.
- Birnbaum, M. H. (1993) *Optometric Management of Nearpoint Vision Disorders*. Butterworth-Heinemann, Boston, pp. 53–71, 161–192.
- Bland, J. and Altman, D. G. (1986) Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* **1**, 307–310.
- Cacho, M. P., García-Muñoz, A., García-Bernabeu, J. R. and López, A. (1999) Comparison between MEM and Nott dynamic retinoscopy. *Optom. Vis. Sci.* **76**, 650–655.
- Campbell, C. E., Benjamin, W. J. and Howland, H. C. (1998) Objective refraction: retinoscopy, autorefraction and photorefracton. In: *Borish's Clinical Refraction* (ed. W. J. Benjamin), Saunders, Philadelphia, pp. 559–628.
- Chat, S. W. and Edwards, M. H. (2001) Clinical evaluation of the Shin-Nippon SRW-5000 autorefractor in children. *Ophthalmic Physiol. Opt.* **21**, 87–100.
- D'augustino, R. B. and Stevens, M. A. (1986). *Tests for the Normal Distribution*. Marcel Dekker, New York.
- Elliott, D. B. (2007). *Clinical Procedures in Primary Eye Care*. Butterworth Heinemann Elsevier, Edinburgh. pp. 194–195.
- Goodson, R. A. and Afanador, A. J. (1974) The accommodative response to the near point crossed cylinder test. *Optom. Wkly.* **65**, 1138–1140.
- Goss, D. (1992) Clinical accommodation testing. *Curr. Opin. Ophthalmol.* **3**, 78–82.
- Goss, D. A., Groppel, P. and Dominguez, L. (2005) Comparison of MEM retinoscopy & Nott retinoscopy & their interexaminer repeatabilities. *J. Behav. Optom.* **16**, 149–155.
- Gwiazda, J., Thorn, F., Bauer, J. and Held, R. (1993) Myopic children show insufficient accommodative response to blur. *Invest. Ophthalmol. Vis. Sci.* **34**, 690–694.
- Kimura, S., Hasebe, S. and Ohtsuki, H. (2007) Systematic measurement errors involved in over-refraction using an autorrefractor (Grand-Seiko WV-500): is measurement of accommodative lag through spectacle lenses valid? *Ophthalmic Physiol. Opt.* **27**, 281–286.
- Levine, S., Ciuffreda, K. J. and Selenow, A. (1985) Clinical assessment of accommodative facility in symptomatic and asymptomatic individuals. *J. Am. Optom. Assoc.* **56**, 286–290.
- Locke, L. C. and Somers, W. (1989) A comparison study of dynamic retinoscopy techniques. *Optom. Vis. Sci.* **66**, 540–544.
- Mallen, E. A. H., Wolffsohn, J. S., Gilmartin, B. and Tsujimura, S. (2001) Clinical evaluation of the Shin-Nippon SRW-5000 autorrefractor in adults. *Ophthalmic Physiol. Opt.* **21**, 101–107.
- McClelland, J. K. and Saunders, K. J. (2003) The repeatability and validity of dynamic retinoscopy in assessing the accommodative response. *Ophthalmic Physiol. Opt.* **23**, 243–250.
- Nakasuka, C., Hasebe, S., Nonaka, F. and Ohtsuki, H. (2003) Accommodative lag under habitual seeing conditions: comparison between adult myopes and emmetropes. *Jpn. J. Ophthalmol.* **47**, 291–298.
- Nakasuka, C., Hasebe, S., Nonaka, F. and Ohtsuki, H. (2005) Accommodative lag under habitual seeing conditions: comparison between myopic and emmetropic children. *Jpn. J. Ophthalmol.* **49**, 189–194.
- Rosenfield, M. and Gilmartin, B. (1990) Effect of target proximity on the open-loop accommodative response. *Optom. Vis. Sci.* **67**, 74–79.
- Rosenfield, M., Portello, J. K., Blustein, G. H. and Jang, C. (1996) Comparison of clinical techniques to assess the near accommodative response. *Optom. Vis. Sci.* **73**, 382–388.
- Rouse, M. W., London, R. and Allen, D. C. (1982) An evaluation of the monocular estimate method of dynamic retinoscopy. *Am. J. Optom. Physiol. Opt.* **59**, 234–239.
- Saladin, J. J. (1998) Phorometry and stereopsis. In: *Clinical Refraction* (ed. W. J. Benjamin), Saunders Company, Philadelphia, pp. 724–773.
- Scheiman, M. and Wick, B. (2008). *Clinical Management of Binocular Vision. Heterophoric Accommodative and Eye movement Disorders*. Lippincott Williams & Wilkins, Philadelphia, pp. 9, 20, 58–76
- Whitefoot, H. and Charman, W. N. (1992) Dynamic retinoscopy and accommodation. *Ophthalmic Physiol. Opt.* **12**, ???–???
- Wick, B. and Hall, P. (1987) Relation among accommodative facility, lag, and amplitude in elementary school children. *Am. J. Optom. Physiol. Opt.* **64**, 593–598.
- Wolffsohn, J. S., Gilmartin, B., Mallen, E. A. and Tsujimura, S. (2001) Continuous recording of accommodation and pupil size using the Shin-Nippon SRW-5000 autorrefractor. *Ophthalmic Physiol. Opt.* **21**, 108–113.
- Wolffsohn, J. S., Gilmartin, B., Mallen, E. A. and Tsujimura, S. (2004) Simultaneous continuous recording of accommodation and pupil size using the modified Shin-Nippon SRW-5000 autorrefractor. *Ophthalmic Physiol. Opt.* **24**, 142–147.
- Zadnik, C. (1997). *The Ocular Examination. Measurements and Findings*. WB Saunders Company, Philadelphia, pp. 98–103.
- Zadnik, K., Mutti, D. O. and Adams, A. J. (1992) The repeatability of measurement of the ocular components. *Invest. Ophthalmol. Vis. Sci.* **33**, 2325–2333.